THE CDER COMMON DATA STANDARDS ISSUES DOCUMENT

AND THE SDTM/ADAM IMPACT

Micaela Salgado-Gomez
Submission Standards and eCTD Submission Strategies

John Brega
PharmaStat LLC

San Diego CDISC Users Group Meeting
03November2011

New Documents

- CDER Common Data Standards Issues Document
  - Issued by CDER to convey
    - Common Issues
    - Requests for future submissions
- Amendment 1 to the SDTM v1.2 and the SDTMIG: Human Clinical Trials v3.1.2 ("Amendment 1")
  - Coordinated effort by CDISC (SDS) to support the requests made by CDER in the Common Data Standards Issues Document
    - Additional variables for many datasets
    - This document is temporarily not available on the CDISC website because the comment period has closed.

Common Issues Document: Version 1.0/May 2011

- In the document CDER,
  - Expresses support for standards
    - References SDTM, SDTMIG, Amendment 1 to SDTM/SDTMIG, SEND, ADaM, ADaMIG, define.xml
  - Encourages sponsor/review division discussions
  - Communicates quality/conformance issues
  - Introduces new standards items
  - Communicates that there will be future updates to the Common Issues Document*

*http://www.fda.gov/Drugs/DevelopmentApprovalProcess/FormsSubmissionRequirements/ElectronicSubmissions/ucm248635

CDERs subtle message that they encourage dialogue with Sponsors

CDER Encourages Discussion With Review Divisions

Page 1 : *"This document is not intended to replace the need for* ***sponsors to communicate with review divisions*** *regarding data standards implementation approaches or issues, but instead, it is designed to complement and facilitate the interaction between sponsors and divisions."*

Page 2 (RE: SDTM) : *"..there may be instances in which the current implementation guides do not provide specific instruction as to how certain study data should be represented. In this instance,* ***sponsors should discuss their proposed solution with the review division*** *and submit supporting documentation as part of the reviewer's guide at the time of submission that describes these decisions/solutions."*

Page 3 (RE: SDTM): *"If there is uncertainty with regards to implementation,* ***the sponsor should discuss with the review division.***"

Page 4 (RE: ADaM): *"It is* ***expected that significant discussion between the sponsor and CDER reviewers will be necessary*** *to appropriately determine which analysis datasets and associated content are needed to support application review."*

Page 4 (RE: SEND): *"****Sponsor should contact the appropriate division*** *with any additional questions."*

Page 7 (RE: Lab Data) : *"..however, it is recommended to* ***confirm this with the review division***."

Questions regarding the content of this document should be submitted to cder-edate@fda.hhs.gov

CDER Feedback – Issues and Requests

Possible Solutions and Their Impact

Issues

- Exact case not used for MedDRA terms
  - Communicate expectation for whoever is responsible for the coding
  - Add review of terms to quality control process
  - Impact: Additional internal/vendor standards that need to be monitored
- Invalid ISO 8601 dates
  - Training
  - Build in processes to create/QC conforming ISO 8601
  - Impact: Time to develop process to create/QC ISO 8601 date/time
- USUBJID not unique across an entire submission for an individual (not EXACT matches)
  - Get information from those who know how to identify these individuals
  - Build in process to identify if these cases exist (likely not programmatic) and if they have been assigned a unique USUBJID across the submission
  - Impact: Need sleuthing time & involvement of other teams for production and review

Issues

- Begin Date must be < = End Date (e.g., AE, CM)
  - Build in process to evaluate dates during production and QC
  - Impact: Time to develop process and to investigate data issues (legacy data)
- Required variables not submitted
  - Build in a process that systematically creates these variables and/or in QC identifies if the variable is not populated
  - Impact: Time to develop process
  - Impact: Time to resolve issues when data is not available
- Inconsistent value for standard units
  - Build in QC process to identify inconsistent values
  - Impact: Time/effort to standardly implement this check in quality control efforts .
  - Impact: Time to correct inconsistent units (legacy data)

QC Help for Some Issues: OpenCDISC

(http://www.opencdisc.org)

"OpenCDISC is an open source community focused on building extensible frameworks and tools for the implementation and advancement of CDISC Standards"

Validator for,

SDTM 3.1.2 , SDTM 3.1.1, Define.xml 1.0, ADaM 1.0, SEND 3.0 soon.

- Since Amendment 1 to the SDTM I.G. 3.1.2 has not been released, OpenCDISC has not been updated to reflect these changes
  - 'False Positives' in OpenCDISC evaluations
  - Need to build in custom checks in the interim

CDER Requests

- Use CDISC Controlled Terminology
  - If it does not exist, propose new terminology
  - Impact: Time to remap values to CT
  - Impact: Time to exhaust terminology and document additions
- Use Consistent Dictionary
  - Impact: Recoding involves additional time and investment
- No imputed data
  - Impact: Save for ADaM
- 1 record per subject in DM
- ADaM Dataset label should not be the same as the label of an SDTM dataset
- Impact: Establish internal/vendor standards and add to quality control process
- If DEATH occurs, it should be the last record in DS and has an EPOCH
  - Impact: Establish internal/vendor standards and add checks to quality control process

CDERs Introduction of New Standards Items and Their Potential Impact

Demographics Domain (DM)

- Strongly preferred added variables

  REQUIRED
  - ACTARMCD/ACTARM: Actual Arm Code /Actual Arm
    - Must be a pre-defined ARMCD/ARM or 'Unplanned' or 'Not Treated'
      - Still no place in domain for Actual Treatment Received if not pre-defined

  EXPECTED
  - RFXENDTC : Date/Time of First Study Treatment Exposure
  - RFENDTC: Date/Time of Last Study Treatment Exposure
  - RFICDTC: Date/Time of Informed Consent
  - RFPENDTC: Date/Time of End of Participation
  - DTHDTC: Date of Death
  - DTHFL: Subject Death Flag

Demographics Domain (DM) cont'd

- Derived Variables
- Increase in development and QC time
  - Additional specifications needed for DM
    - Does the protocol/SAP define these variables well enough for programmers/vendor to derive?
    - Study specific
    - Additional programming (beyond database capabilities)
      - Data issues (Missing/partial/'bad' dates/legacy data)
    - Cross reference quality concerns
      - Same data exists in DS/DM domain
        - RFPENDTC: Date/Time of End of Participation
        - DTHDTC: Date of Death
        - RFICDTC: Date/Time of Informed Consent
        - RFSTDTC / RFENDTC
- Change to Process
  - Is this an appropriate task for those producing SDTM?
  - Allow time, assign tasks to verify that variables have been applied correctly as defined by sponsor and 'make sense' for the study.

ADaM Impact: Demographics

- Per Subject
  - Key variables in a standard location that don't need to be re-derived if QCd (specifications are accurate, output is consistent with specifications, and values 'make sense')

Adverse Events Domain (AE)

- AE domain should contain all adverse events recorded for subjects
  - Sponsor should not filter based on their evaluation of 'treatment emergent'
  - Flag sponsor defined treatment emergent events (as used in the sponsor's primary adverse events analysis) by using the variable, AETRTEM*
- Increase in development and QC time
  - Additional specifications needed for AE (Study Specific)
    - Does the protocol/SAP define treatment emergent well enough for programmers/vendor to derive?
    - Study specific.
    - Additional programming (beyond database capabilities)
      - Data issues (Missing/partial/'bad' dates/legacy data)
- Change to Process
    - Is this an appropriate task for those producing SDTM?
    - Allow time, assign task to verify SDTM treatment emergent flag has been applied correctly as defined by sponsor and to ensure that SDTM 'treatment emergent is the same as treatment emergent used in primary AE analysis

* The –TRTEM variable may also be used for other Events datasets per draft Amendment 1 to SDTM v1.2 and SDTMIG v3.1.2

Adverse Events Domain (AE)

- 10 new EXPECTED variables for MedDRA terms *
  - MedDRA terms other than SOC and preferred term previously reported in SUPP– when coding existed
  - IG and CDER at odds about whether primary or secondary SOC should be used in AEBODSYS (IG allows primary or secondary to be used to populate AEBODSYS)
    - Need to allow reviewers to easily determine whether the use of a secondary SOC for analysis was appropriate
      - Dedicate AESOC for MedDRA defined primary mapped SOC
      - Dedicate AEBODSYS for the secondary mapped SOC

  <u>Change to process</u>
  - Create variables regardless of whether the variables are populated
  - Populate AEBODSYS the same from study to study

*These variables may also be used for other *Events* datasets per draft Amendment 1 to SDTM v1.2 and SDTMIG v3.1.2

ADaM Impact: Adverse Events

- Re-train to use AESOC/AEBODSYS in production/QC
- Treatment emergent should not be re-derived. Definition for SDTM and ADaM should be the same.

Disposition Domain (DS)

- When there is more than one disposition event, the EPOCH variable should be used to distinguish between them

Increase in development and QC time

- Study specific. Needs to be defined before development
- Ensure consistent with Trial Design
- In QC, identify whether more than one disposition event exists to ensure that EPOCH has been populated

ADaM Impact: Disposition

- Appropriate disposition should be consistently easier to identify using the EPOCH variable.

Supplemental Qualifiers (SUPP--)

- Not a "waste basket" for data elements that the sponsor is not sure how to allocate.
  - Requires being more conscientious with variables that don't "fit" into the existing models
    - Is your data (reasonably) usable for a reviewer?
    - Consider another dataset /custom domain
    - Does variable need to be submitted?
  - Key analysis variables
    - Discuss with reviewer if key analysis variables will be included in SUPP
    - If needed for ADaM, keep for traceability
    - Document
  - Extraneous variables
    - Science and regulation should drive what is submitted

ADaM Impact: Supplemental Qualifiers

- Source of key variables may be found in supplemental qualifiers

Subject Level Data: Add EPOCH, ELEMENT, ETCD

- Add variables to allow reviewer to determine during which phase of the trial the observation occurred as well as actual intervention the subject experienced during that phase

  - Highly Derived Variables
    - **Increased development and QC time**
      - Additional specifications needed for SDTM
        - Does the protocol /SAP define the epochs well enough for programmers/vendor to derive? Study specific.
      - Additional programming (beyond database capabilities)
        - Data issues (Missing/partial/'bad' dates)
        - Determining EPOCH may be difficult if defined by events rather than visits, and involves ***date windowing*** for log data like AE and CM
      - Additional review required
        - Statistician/sponsor involvement to ensure that specifications are correct and results are as expected ('make sense')
    - **Change to process**
      - Appropriate task for those currently responsible for producing SDTM?
      - Add review to timeline to 'work out the details'
      - Where do boundary dates get stored? Use a macro? Submitted?

SDTM Dataset Size

- Some files are too big:
  - Recommend submitting smaller datasets broken up by a meaningful variable like –CAT. Submit both smaller dataset and the cumulative dataset.
    - Process

      Breaking up the data
      - "Large" is not really defined ("400 megabytes is usually fine")
        - Guess what? *"confirm this with the review division"*
      - Will all datasets be evaluated for size?
      - What variable will be used to break up datasets (Case by case?)

      Describing data in the define.xml
      - No recommendation about how this should be dealt with in the define.
        - Include 2x since the data is submitted 2x ?
        - Include an explanation in the tabulations data guide?
    - Quality Control
      - Ensure that parts make up the whole
      - Ensure that attributes are the same across the parts
      - Ensure that documentation is clear about the data

ADaM Impact: SDTM Dataset Size

- Use large SDTM dataset to produce ADaM dataset?
- If after processing, data is still large, separate data again to more reasonable size?
  - Same quality concerns as SDTM
- Traceability? Back to cumulative SDTM dataset?

SDTM Variables : General

- Permissible does not mean optional
  - "All permissible variables for which data were collected or for which derivations are possible should be submitted"
    - Examples:
      - Baseline flags for LB, VS, EG, PC, MB
      - EPOCH designators
      - --DY and –-STDY (calculated based on first treatment day) variables in Findings domains or SE
    - Impact
      - More variables, more programming and QC

SDTM Variables: Date/Time

- Always ISO8601
- Do not pad with trailing zeros

ADaM Impact: SDTM Variables

- Baseline flags
  - Inherited from SDTM?
  - Different (more complicated) derivations for analysis than for SDTM?
- Variable definition consistency
  - --DY and –-STDY now calculated based on first treatment day for Findings domains
- Date/time precision is accurate
  - Allowing for more direct date/time elapsed calculation

Acknowledgments

Jane Diefenbach
PharmaStat LLC

Thank you

Questions?

Micky Salgado-Gomez
Submission Standards and eCTD Strategies
mickygomez@gmail.com